Journal of Hazardous Materials 213–214 (2012) 83–92

Contents lists available at SciVerse ScienceDirect

Journal of Hazardous Materials

journal homepage: www.elsevier.com/locate/jhazmat

Prediction of the wash-off of traffic related semi- and non-volatile organic compounds from urban roads under climate change influenced rainfall characteristics

Parvez Mahbub[a,*], Ashantha Goonetilleke[a], Godwin A. Ayoko[b]

[a] School of Urban Development, Queensland University of Technology, GPO Box 2434, Brisbane 4001, Queensland, Australia
[b] Chemistry Discipline, Queensland University of Technology, GPO Box 2434, Brisbane 4001, Queensland, Australia

ARTICLE INFO

Article history:
Received 14 May 2011
Received in revised form 17 January 2012
Accepted 19 January 2012
Available online 25 January 2012

Keywords:
Semi-volatile organic compounds
Non-volatile organic compounds
Pollutant wash-off
Climate change

ABSTRACT

Traffic generated semi- and non-volatile organic compounds (SVOCs and NVOCs) pose a serious threat to human and ecosystem health when washed off into receiving water bodies by stormwater. Climate change influenced rainfall characteristics makes the estimation of these pollutants in stormwater quite complex. The research study discussed in the paper developed a prediction framework for such pollutants under the dynamic influence of climate change on rainfall characteristics. It was established through principal component analysis (PCA) that the intensity and durations of low to moderate rain events induced by climate change mainly affect the wash-off of SVOCs and NVOCs from urban roads. The study outcomes were able to overcome the limitations of stringent laboratory preparation of calibration matrices by extracting uncorrelated underlying factors in the data matrices through systematic application of PCA and factor analysis (FA). Based on the initial findings from PCA and FA, the framework incorporated orthogonal rotatable central composite experimental design to set up calibration matrices and partial least square regression to identify significant variables in predicting the target SVOCs and NVOCs in four particulate fractions ranging from >300 to 1 μm and one dissolved fraction of <1 μm. For the particulate fractions in >300–1 μm range, similar distributions of predicted and observed concentrations of the target compounds from minimum to 75th percentile were achieved. The inter-event coefficient of variations for particulate fractions of >300–1 μm was 5–25%. The limited solubility of the target compounds in stormwater restricted the predictive capacity of the proposed method for the dissolved fraction of <1 μm.

1. Introduction

Traffic related semi- and non-volatile organic compounds (SVOCs and NVOCs) are primarily associated with diesel fuels, fuel oils, heavier engine oils and lubricants [1]. Homologous series of n-alkanes from decane to tetracontane are amongst the most common constituents of these products, which are widely used in motor vehicles, and have the potential to pollute the urban water environment through deposition and wash-off from urban roads [2]. Rainfall characteristics such as, intensity, duration and frequency or average recurrence intervals (ARIs) are predicted to undergo significant changes as a result of climate change. In this context, the Commonwealth Scientific and Industrial Research Organisation (CSIRO) has forecasted longer periods of dry weather with fewer, but more intense storms in Australia due to climate change [3]. Such climate change driven changes in the rainfall characteristics will affect the wash-off processes of various stormwater pollutants including the SVOCs and NVOCs.

The detrimental effects of SVOCs and NVOCs on human health have been widely reported in research literature. Mutagenic evidence in mammalian cells caused by diesel engine exhaust particles has been cited by Bao et al. [4]. Morgan et al. [5] attributed the long term exposure to diesel engine exhaust particles to respiratory allergy, cardiopulmonary mortality and risk of lung cancer. Petroleum related activities have also been attributed to significant wetland loss in the Mississippi Delta [6]. While studies on the impacts of traffic generated volatile organic compounds such as BTEXs (benzene, toluene, ethylbenzene and xylene) in urban roads [7] and ambient atmosphere have commonly been undertaken [8,9], such pollutants have only been characterised in terms of concentrations and modelled for the ambient atmosphere [10]. However, it is important to note that pollutants present in the urban atmosphere are not necessarily deposited on the urban roads due to various climatic factors. Therefore, compartment-based multimedia models (e.g. separate wash-off models from pervious and

* Corresponding author. Tel.: +61 7 3138 1092; fax: +61 7 3138 1170.
E-mail addresses: s.mahbub@qut.edu.au (P. Mahbub), a.goonetilleke@qut.edu.au (A. Goonetilleke), g.ayoko@qut.edu.au (G.A. Ayoko).

0304-3894/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhazmat.2012.01.062

impervious surfaces) are particularly suitable in rationalising the differences in environmental fate and transport of pollutants in a defined environment [11]. Nevertheless, the wash-off phenomenon from urban roads becomes complex when the changed rainfall characteristics due to climate change affect the wash-off processes of such pollutants. In this context, the current state of knowledge on traffic generated semi- and non-volatile organic compounds (SVOCs and NVOCs) available on roads for wash-off is very limited.

Mahbub et al. [12] recently proposed a prediction model for the build-up of five traffic generated volatile organic compounds on urban roads. However, their model did not investigate the uncertainties involved in the wash-off of the wide range of traffic generated pollutants from roads under climate change driven changes to rainfall characteristics. Accurate estimations of the concentrations of available SVOCs and NVOCs on roads in wash-off under climate change are required in order to undertake mitigation measures for the management of such pollutants in stormwater runoff. Accordingly, this research study presents a framework for predicting the concentrations of traffic generated SVOCs and NVOCs in wash-off under climate change influenced rainfall characteristics. This approach is expected to contribute to overcome the uncertainties inherent in the wash-off estimation of traffic generated SVOCs and NVOCs by predicting these pollutants based on the significance of individual predictors and consequently, strengthening the appropriate measures for pollution mitigation.

2. Materials and methods

2.1. Site selection

Four road sites within a 5 km radius from a meteorological gauging station were selected as the wash-off study sites. The station was located at 27.90°S and 153.31°E at an elevation of 6 m above mean sea level with daily rainfall data recorded since 1894. The selected road sites were situated in three relatively new suburbs in the Gold Coast region, Australia with the transport infrastructure developed in the last decade. The sites were in different land uses such as residential, commercial and industrial in order to incorporate a mix of vehicular traffic characteristics. The locations, traffic and pavement characteristics of the selected sites are provided in the supplementary data. Due to their close proximity to the rain gauging station, it was hypothesised in this study that the predicted changes in the rainfall characteristics at the four study sites resulting from climate change are similar to that at the rain gauging station.

2.2. Rainfall simulation incorporating climate change

The research study used a rainfall simulator [13] to replicate the design rainfall events resulting from climate change. The rainfall simulation was based on the studies of Abbs et al. [14] who predicted the average fractional change for extreme rainfall intensities at 2, 24 and 72 h durations for the Gold Coast area in Australia for 2030 and 2070 using CSIRO general circulation model known as CC-MK3 and CSIRO regional downscaling model known as RAMS. Several climate change studies [3,15] have predicted that the probability of occurrence of shorter duration (<2 h) events with a large change in precipitation intensities is very high.

Mahbub et al. [16] used the outcome from the Abbs et al. [14] study and proposed the following three scenarios to describe the climate change influenced rainfall characteristics in the Gold Coast region:

- Shorter duration, with higher intensity with ARI constant;
- Shorter ARI, shorter duration with intensity constant; and
- Shorter ARI, with higher intensity while duration becomes shorter.

The current study incorporated these scenarios by simulating the 2009 and 2030 rainfall characteristics in the Gold Coast region of Australia according to the study by Mahbub et al. [16]. As the subsequent chemometric data analyses and interpretations require referencing to these simulated rainfall events, Table 1 is reproduced in this paper.

A total of twenty-two rain events were simulated in the four selected road sites. It was not feasible to simulate all twenty-two rain events simultaneously at all four sites due to time restrictions imposed by the city council on road lane closures. Therefore, the simulation events were distributed among the four study sites in different sets of intensity ranges of 24.6–39.3, 58.3–63, 75–77 and 119–125 mm/h.

2.3. Wash-off sample collection

The rainfall simulations were undertaken over a 2 month period from April to May 2009. Wash-off samples resulting from the simulations were collected using a commercially available vacuum cleaner. The weather was dry and the temperature during the sampling ranged between 22 °C and 25 °C. The collection plots were 3 m^2 in size and were located in the middle of the traffic lanes at the study sites, marked with permanent markers, and thoroughly cleaned with deionised water. Then the plots were left for seven dry days to allow for traffic generated pollutants to build-up. This allowance of seven dry days was in conformity with the findings of Egodawatta [17] who noted that the pollutant build-up on road surfaces asymptote to an almost constant value after an antecedent dry period of 7 days. The collection plots were connected to a collection trough [13]. The runoff water in the collection trough was vacuumed continuously into 25 L plastic containers. The plastic containers were washed thoroughly inside out with 10% HCl followed by Decon 90® detergent wash and rinsed throughout with deionised water. The containers were then dried at 40 °C for 48 h before collecting samples from the field. A photo of the sample collection procedure is provided in the supplementary data.

After collection, the runoff samples were transported to the laboratory for sub-sampling immediately. As pollutant concentrations can vary by orders of magnitude during a runoff event, the flow weighted average or event mean concentration (EMC) samples were found to be appropriate for evaluating the impacts of stormwater runoff on receiving waters [18]. In this study, 500 mL EMC samples in amber glass bottles were prepared in the laboratory using a churn splitter. The required volumes at a particular duration constituting an EMC sample were determined from the percentages of the total runoff collected in different containers for that duration and mixed together to obtain the 500 mL EMC sample for an event.

The particle size distributions of the suspended solids in the subsamples were determined using a Malvern Mastersizer S Particle Size Analyser capable of analysing particles between 0.05 and 900 μm diameter. The particle size distributions of the subsamples were used as a guide for maintaining homogeneity in the sub-samples throughout the sample splitting process. Based on the particle size distribution, the total particulate analytes in the 500 mL EMC subsamples were fractioned into four size ranges, namely, >300 μm, 150–300 μm, 75–150 μm, 1–75 μm using wet sieving. The filtrate passing through a 1 μm membrane filter was considered as the total dissolved fraction. In each case, 500 mL homogeneous sub-samples were prepared using deionised water, collected in 500 mL amber glass bottles with a PTFE seal, preserved with 5 mL of 50% HCl at 4 °C in the laboratory and analysed within 40 days of collection. A total of 110 wash-off samples were

Table 1
Simulation events based on the daily rainfall intensity at study sites in the Gold Coast region for 2030.

Scenario	Simulation events for Gold Coast region for 2009				Future simulation events for Gold Coast region for 2030			
	Simulation event	Duration (min)	Intensity (mm/h)	ARI (year)	Simulation event	Duration (min)	Intensity (mm/h)	ARI (year)
Shorter duration, with higher intensity with ARI constant	1	60	39.3	1	19	25	63	1
	3	90	39.3	2	20	42.5	61.2	2
	5	133	39.3	5	21	69	59.2	5
	6	160	39.3	10	22	85	58.3	10
	18	105	75	100	13	49	115	100
		–	–	–	2	65	37.39	1
Shorter ARI, shorter duration with intensity constant	12	45	125	100	7	5	125	1
		–	–	–	4	120	24.6	1
Shorter ARI, with Higher intensity whilst duration becomes shorter	14	52.5	77	10	10	16	125	5
	15	67.5	77	20	11	21	122	10
	16	86.7	77	50	9	10.5	120	2
	17	101.25	77	100	8	5.75	119	1

Adapted from Mahbub et al. [16].[a]

[a] Copyright: IGI Global Source Publication; reprinted by permission of the publisher.

prepared for the 22 simulated rain events with each event consisting of five samples based on the size fractions mentioned above.

The extractions of SVOCs and NVOCs were performed by liquid–liquid extraction with 250 mL hexane as the exchange solvent according to USEPA method 3510C [19]. The extracted samples were then cleaned using standard column cleanup protocol with 5 cm silica gel and 5 cm pyrex® glass wool topped with 5 cm anhydrous Na_2SO_4 [19]. Sample concentration was then carried out using the Kuderna–Danish apparatus followed by the nitrogen blowdown technique [19]. The sample concentration was continued until a final volume of 1 mL was achieved for gas chromatographic (GC) analyses.

2.4. Sample testing

Based on USEPA approved methods for the determination of diesel range organics, Wisconsin DNR [20] has identified seventeen traffic generated SVOCs and NVOCs (from octane to tetracontane) as constituents from the larger organic compound group (i.e. diesel range organics) and developed modified methods for their determination in soil, stormwater and wastewater matrices. Whilst, both petrol and diesel engine vehicles emit gaseous and particulate hydrocarbons as a result of incomplete combustion [21], Andreou and Rapsomanikis [22] noted that past studies mainly characterised only one organic group (polycyclic aromatic hydrocarbons). The mutagenic and other detrimental impacts of traffic generated SVOCs and NVOCs on human health has been cited in research literature (for example 4–6). Therefore, this research study focused on the above mentioned seventeen traffic generated SVOCs and NVOCs.

The target SVOCs for the study were octane (OCT), decane (DEC), dodecane (DOD), tetradecane (TED), hexadecane (HXD), octadecane (OCD), eicosane (EIC), docosane (DOC), tetracosane (TTC), hexacosane (HXC), and octacosane (OCC) having boiling points ranging from 125 °C to 432 °C [23]. For the convenience of the predictive framework proposed in the study, the target SVOCs were further separated into two groups based on their molecular weights, namely 'light SVOC' and 'heavy SVOC'. The 'light SVOC' group consisted of four SVOCs from octane to tetradecane whilst the 'heavy SVOC' group consisted of the remaining seven SVOCs from hexadecane to octacosane. The target NVOCs were triacontane (TCT), dotriacontane (DTT), tetratriacontane (TRT), hexatriacontane (HXT), octatriacontane (OTT), and tetracontane (TTT) with boiling points ranging from 449 °C to 525 °C [23].

USEPA methods 3510C, 8015, 8021, and 8260 [19] were adopted for the determination of SVOCs. Draper et al. [2] proposed modifications to the USEPA methods to determine motor oils with carbon numbers up to C_{38}. This study used these modifications as a guide to establishing the gas chromatographic (GC) temperature program for simultaneous determination of both SVOCs and NVOCs. Details of SVOC and NVOC test methods, chemical compositions as well as relative comparisons of chemical concentrations with past studies are provided in the supplementary data.

Other physico-chemical variables such as total suspended solid (TSS) and total organic carbon (TOC) were determined by methods 2540D and 5310B [24]. Additionally, the pH and electrical conductivity (EC) of each sample were measured using standard pH and EC probes in the laboratory according to methods 4500-H^+ B and 2510B, respectively [24].

2.5. Data analysis

Data matrices were constructed for light SVOCs, heavy SVOCs and NVOCs for the five size fractions noted above. Each matrix consisted of twenty-two objects with numerical object identifiers (same as the simulation events in Table 1) starting with 1. Rainfall characteristics such as, intensity, frequency, and duration as well as the physico-chemical characteristics such as TSS, TOC, pH, and EC were considered to be the independent variables causing the wash-off of the target SVOCs and NVOCs. After initial observation of the probability distribution of the objects and variables, standardisation of each variable and normalisation of each object were undertaken as pre-treatment measures.

The data analysis was designed to investigate the wash-off process of SVOCs and NVOCs under climate change conditions and then to apply the findings from the initial investigations to develop a prediction framework for light SVOCs, heavy SVOCs and NVOCs wash-off. Multivariate chemometrics methods such as principal component analysis (PCA), factor analysis (FA), experimental design, and partial least squares (PLSs) regression were employed for the data analysis. Discussions of these techniques are given in the supplementary data.

3. Results and discussion

3.1. Exploratory principal component analysis

Wash-off data matrices for light SVOCs, heavy SVOCs and NVOCs were analysed for all five size fractions. Fig. 1 shows the PCA biplots for total particulate (<300–1 μm) and dissolved fractions (<1 μm). This study adopted the rain events classification under climate change proposed by Mahbub et al. [25]. Events with intensity <40 mm/h with relatively low ARI were classified as low events; those having intensity between 50 and 100 mm/h but with

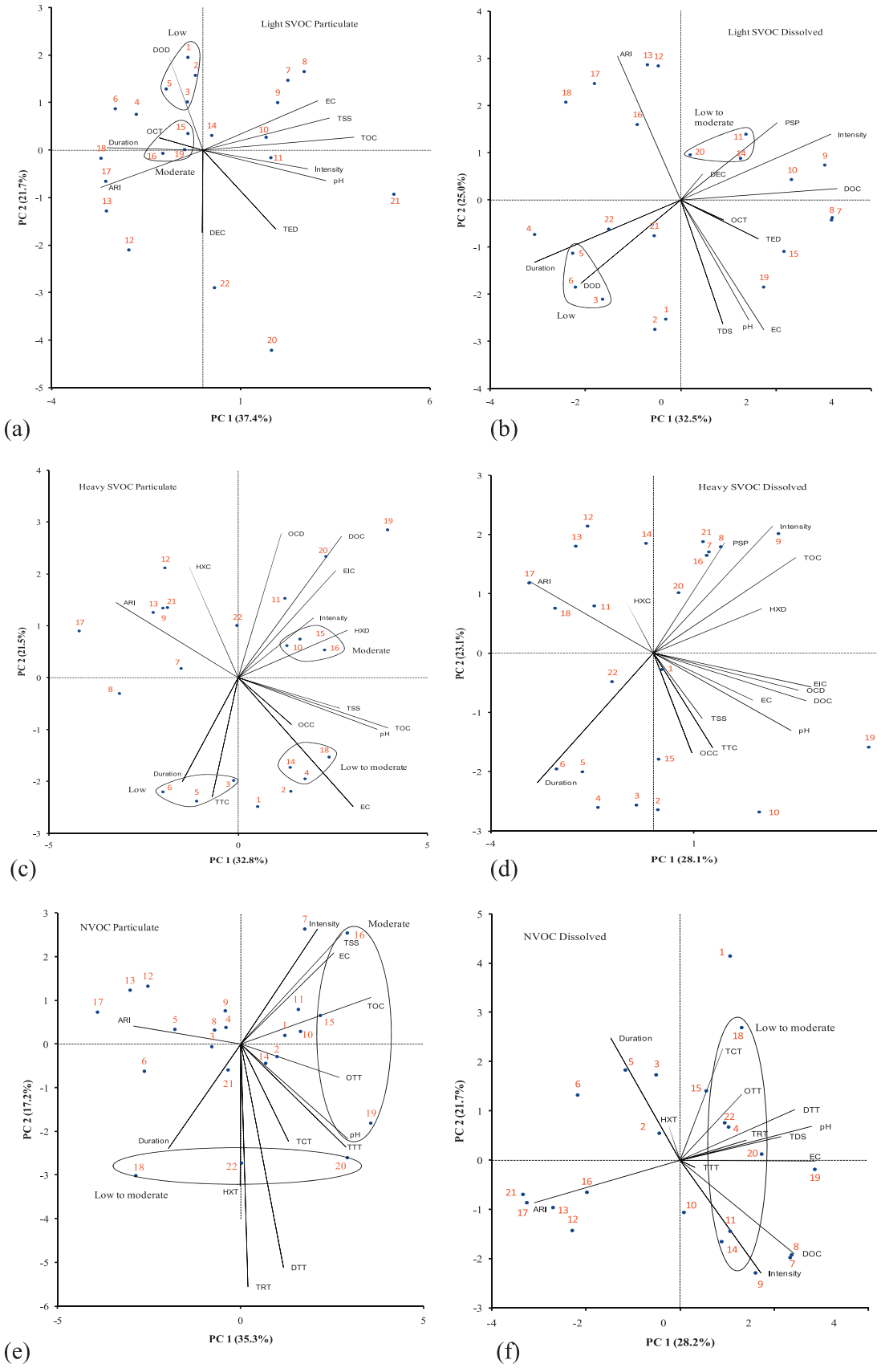

Fig. 1. PCA biplots of particulate (>300–1 μm combined) and the dissolved (<1 μm) fractions for light SVOCs, heavy SVOCs and NVOCs for the 22 rain events shown with numerical identifiers which are same as the simulation event numbers described in Table 1.

Table 2
New independent variables for underlying factors (starting with initials L, H or N) in the data matrices of light SVOC, heavy SVOC, and NVOC.

Rain events	Underlying factors												
	Light SVOC				Heavy SVOC				NVOC				
	L1	L2	L3	L4	H1	H2	H3	H4	N1	N2	N3	N4	N5
1	−1.258	−0.061	1.173	0.195	−0.522	−0.840	−0.232	1.037	−1.194	1.598	−0.342	−0.966	0.708
2	−1.070	0.043	0.987	0.516	−0.668	−0.881	−0.278	0.951	−1.037	1.671	−0.148	0.088	0.423
3	−0.871	−0.554	0.376	0.083	−0.471	−1.034	0.618	0.897	−1.141	0.448	−0.498	1.109	−0.976
4	−1.475	−0.262	0.527	0.257	−0.545	−1.355	0.229	0.704	−1.516	0.575	0.004	0.052	0.172
5	−1.245	−0.412	−0.192	0.366	−0.958	−1.294	−0.119	0.086	−1.423	−0.286	−1.154	−0.337	−0.293
6	−1.701	−0.519	−0.308	−0.281	−0.913	−1.610	−0.340	0.048	−1.510	−0.776	−0.508	−0.580	−1.358
7	1.466	−0.443	2.265	−0.071	−0.747	1.708	−0.008	2.027	1.232	1.651	−0.777	0.583	−1.126
8	1.234	−0.517	2.166	−0.605	−0.276	1.554	0.600	1.981	1.361	1.877	−0.233	0.151	−1.012
9	1.052	0.138	0.390	0.974	0.562	1.165	−0.448	0.235	1.363	0.205	0.206	−0.724	−0.075
10	1.128	−0.120	0.078	−0.052	−0.510	1.137	−0.757	−0.072	1.233	−0.088	0.083	−0.162	−0.439
11	1.377	0.278	−0.910	−0.267	−0.251	1.052	−1.039	−0.608	1.262	−0.564	−0.457	−0.513	0.479
12	1.113	0.326	−1.384	−0.469	0.042	0.818	−1.213	−1.117	0.827	−1.148	−0.683	0.796	−0.209
13	0.978	−0.004	−0.949	3.348	0.008	0.303	−0.668	−0.678	0.419	−1.029	−0.550	1.998	−0.894
14	0.293	−0.677	−0.647	−0.816	−0.628	0.696	0.891	−1.411	0.488	−0.712	−0.427	−1.516	0.710
15	0.212	−0.974	−0.558	−0.892	0.548	0.322	2.140	−0.676	0.147	−0.540	−0.049	−0.900	0.338
16	0.144	−0.845	−0.650	−0.875	0.741	0.035	1.415	−0.382	0.383	−1.026	−0.825	−0.495	1.105
17	0.062	−0.972	−1.120	−0.806	−0.521	0.300	1.013	−1.638	−0.141	−1.157	−0.166	−0.876	−1.051
18	−0.255	−1.256	−0.784	−0.311	−0.240	−0.170	1.711	−0.630	−0.342	−0.552	3.688	−0.236	−1.589
19	−0.074	1.602	0.773	−0.684	3.162	−0.019	−0.369	0.696	0.522	0.961	1.208	0.980	1.171
20	−0.163	3.101	−0.147	−0.862	0.124	−0.158	−1.362	−0.606	0.121	0.190	1.121	−0.452	2.076
21	−0.322	1.516	−0.892	−0.519	2.260	−1.227	−0.095	0.151	−0.590	−0.922	0.270	2.618	1.630
22	−0.625	0.612	−0.194	1.771	−0.197	−0.504	−1.686	−0.996	−0.463	−0.376	0.237	−0.621	0.206

relatively higher ARIs of up to 50 years were classified as moderate events; events having intensities >100 mm/h with very high frequency were classified as high events whilst events with similar intensities to moderate and high with extremely rare occurrence (ARI ≥ 100 years) were classified as extreme events. Events which manifested the attributes of both low and moderate events were classified as low to moderate events.

Based on the classification, events 1, 2, 3, 4, 5, 6 were grouped as low events; events 19, 20, 21 were grouped as low to moderate events; events 14, 15, 16, 22 were grouped as moderate events; 7, 8, 9, 10, 11 were grouped as high events; 12, 13, 17, 18 were grouped as extreme events. In Fig. 1, two important facts were noted. Firstly, the average recurrence intervals (ARIs) are uncorrelated to both the intensities and durations of events as the loading vector of ARI is nearly perpendicular to those of intensities and durations (Fig. 1b–f). Therefore, any prediction framework for SVOCs and NVOCs should not include all three of them together as measured variables. As the intensities and durations were more strongly correlated with the target variables than ARI (Fig. 1a–f), the analysis excluded ARI from the measured variables list in the subsequent analysis. The relative importance of other variables such as, pH, EC, TSS, TOC in the prediction of the target compounds during wash-off was substantiated based on their positive correlations with these compounds in Fig. 1a–f.

The second important fact evident in the biplots of Fig. 1 was that the low, low to moderate and moderate rain events formed clusters strongly correlated to the target compounds during wash-off except in Fig. 1d. This suggested that the low, low to moderate and moderate rain events primarily caused the wash-off of the light SVOCs, heavy SVOCs and NVOCs. These preliminary findings were useful in selecting the experiments (i.e. rain events) to construct the calibration matrices in the experimental design.

3.2. Factor analysis

Factor analysis in two phases, namely, factor extraction and orthogonal varimax rotation was performed to identify the underlying independent factors of the data matrices for light SVOCs, heavy SVOCs and NVOCs. After careful investigation of the rotated component matrices for light SVOCs, heavy SVOCs and NVOCs which consisted of the correlations between the measured variables and the factors, four underlying factors were found to be sufficient for the light SVOC and heavy SVOC matrices whilst five factors were deemed necessary for the NVOC matrix. These independent factors were extracted based on the initial eigenvalue criteria ≥1. The underlying factors were assigned with numerical identifiers each starting from 1 with initials 'L', 'H' and 'N' for light SVOCs, heavy SVOCs and NVOCs respectively. New variables for each factor corresponding to the twenty-two rain events were then created by the regression method [26] as shown in Table 2.

The new variables (i.e. factor scores) generated in Table 2 were used in the subsequent PLS regression models to predict the corresponding target variables.

3.3. Experimental design

Three calibration sets for the PLS model were optimised with two level orthogonal rotatable central composite design for light SVOCs, heavy SVOCs, and NVOCs. As the number of factor levels and their values were unknown in the design, the study incorporated the Sirius software [27] generated coded values for the two levels, namely, high and low and incorporated 35 experiments (28 individual experiments and 7 replicate experiments at centre) for the light SVOCs and heavy SVOC data matrices. Similarly, the study incorporated 50 experiments (46 individual experiments and 4 central replicate experiments) for the NVOC data matrix. A higher number of experiments for NVOCs were required due to the large number of underlying factors in the NVOC data matrix. Experiments were only chosen from low, low to moderate and moderate rain events as these were found to be the primary events causing the wash-off of the target compounds. It was ensured that each of the five size fractions contributed to the calibration matrices by selecting at least seven experiments from each fraction. In Fig. 2, PCA biplots for three calibration sets are shown.

With few exceptions, in Fig. 2a–c, most of the central experiments were found close to the origin of the biplots, which meant that these were replicates of the same or similar experiments and did not need to be included in the design. The central or replicate

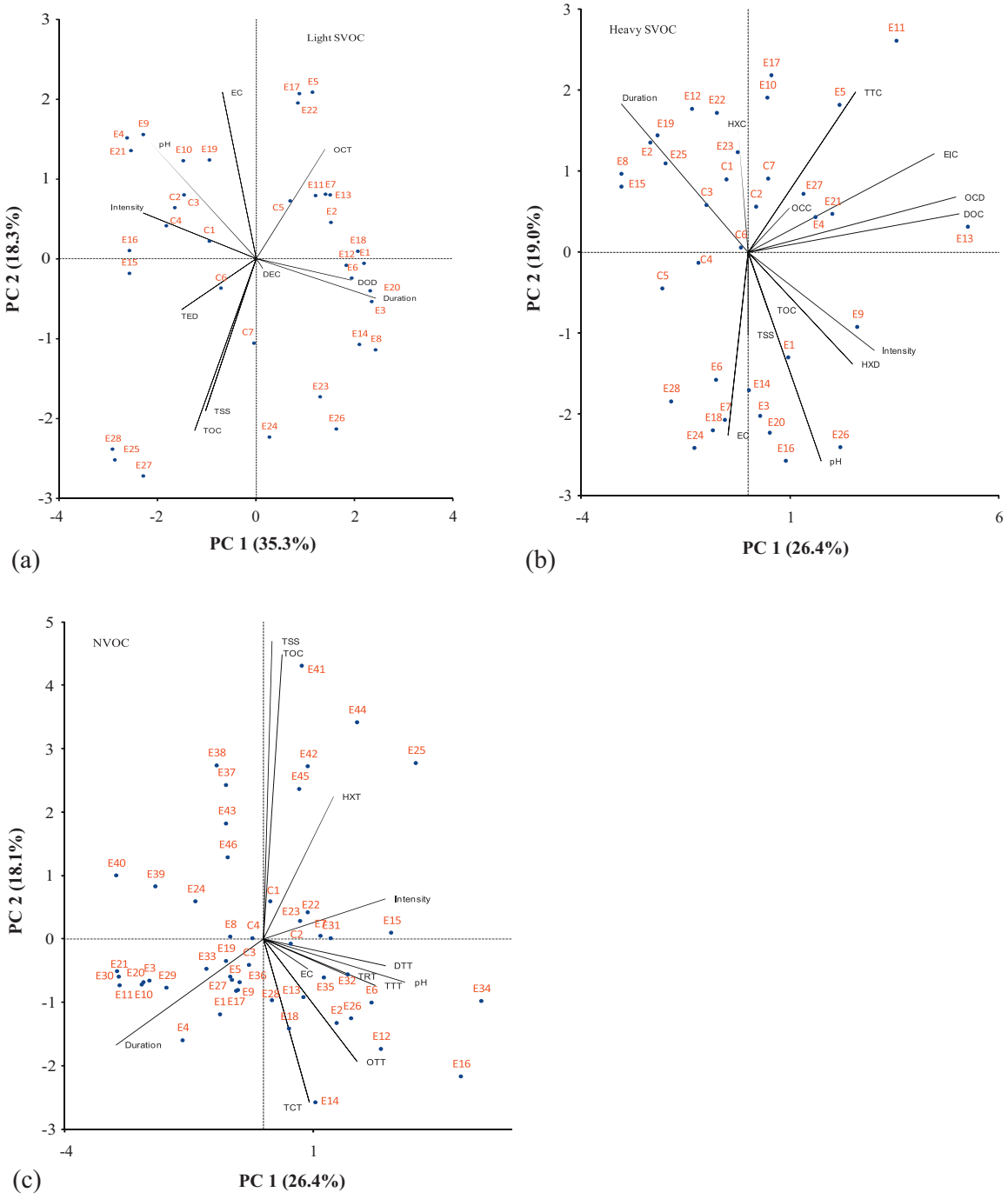

Fig. 2. PCA biplots of the experimental designs for (a) light SVOCs, (b) heavy SVOCs and (c) NVOCs with original experiments are shown with initial 'E' and replicate experiments with initial 'C'.

experiments were chosen in order to identify any curvature present on the response surface by comparing their mean values with that of the rest of the experiments. In Fig. 2a, 12 experiments were found to be very strongly correlated with target the compounds dodecane (DOD) and octane (OCT), in Fig. 2b, 13 experiments were found to be strongly correlated with all target heavy SVOCs whilst in Fig. 2c, 19 experiments were found to be strongly correlated with all target NVOCs. This suggested that the calibration matrices closely corresponded with the wash-off of the target compounds under climate change influenced rainfall characteristics even though the total variances explained by the PCs in Fig. 2 were around 45–53%. The calibration sets are provided in the supplementary data.

3.4. PLS model validation

In the PLS regression, the target compounds OCT, DEC, DOD, TED, HXD, OCD, EIC, DOC, TTC, HXC, OCC, TCT, DTT, TRT, HXT, OTT, and TTT were considered as dependent or measured variables whilst the factors extracted in the factor analysis process along with intensity, duration, TSS, TOC, pH, and EC were considered as the predictor variables. A cross-validation method [28] that left one experiment out at a time from the calibration set was used to measure the standard error in cross-validation (SECV). The following three criteria were employed to determine the required number of PLS components for regression:

Table 3
PLS regression parameters for predictor variables.

Measured variables	PLS components	Variance explained by predictor variables, %	Variance explained by measured variables, %	Coefficient of determination, r^2	SECV	Regression coefficients for predictor variables									
						TSS	TOC	pH	EC	Intensity	Duration	Underlying factors			
												L1[a]	L2[a]	L3[a]	L4[a]
OCT	1	31.87	39.40	0.50	0.94	−0.29	−0.28	−0.18	0.22	I.F.[d]	I.F.[d]	−0.21	I.F.[d]	I.F.[d]	I.F.[d]
DEC	1	65.51	57.55	0.57	0.71	I.F.[d]	−0.03	−0.04	I.F.[d]	−0.04	0.03	I.F.[d]	I.F.[d]	I.F.[d]	I.F.[d]
DOD	1	46.28	45.17	0.50	0.86	I.F.[d]	−0.09	−0.27	0.01	−0.27	0.20	−0.06	I.F.[d]	I.F.[d]	I.F.[d]
TED	1	50.96	46.48	0.50	0.96	I.F.[d]	0.26	0.29	I.F.[d]	I.F.[d]	−0.24	I.F.[d]	0.13	0.03	I.F.[d]

Measured variables	PLS components	Variance explained by predictor variables, %	Variance explained by measured variables, %	Coefficient of determination, r^2	SECV	Regression coefficients for predictor variables									
						TSS	TOC	pH	EC	Intensity	Duration	Underlying factors			
												H1[b]	H2[b]	H3[b]	H4[b]
HXD	1	55.24	51.08	0.60	0.95	0.04	I.F.[d]	0.50	I.F.[d]	I.F.[d]	I.F.[d]	I.F.[d]	I.F.[d]	0.29	I.F.[d]
OCD	1	61.50	57.07	0.67	1.00	I.F.[d]	I.F.[d]	0.08	−0.19	0.22	−0.19	−0.22	−0.13	I.F.[d]	I.F.[d]
EIC	1	52.96	45.03	0.60	1.00	−0.20	−0.10	0.15	−0.23	I.F.[d]	I.F.[d]	−0.39	I.F.[d]	I.F.[d]	I.F.[d]
DOC	1	47.89	42.61	0.63	0.97	I.F.[d]	I.F.[d]	0.20	−0.19	0.47	I.F.[d]	I.F.[d]	−0.23	I.F.[d]	I.F.[d]
TTC	1	55.33	54.25	0.67	1.00	I.F.[d]	I.F.[d]	I.F.[d]	I.F.[d]	−0.21	I.F.[d]	−0.40	I.F.[d]	−0.36	I.F.[d]
HXC	1	46.82	47.64	0.70	0.96	−0.13	−0.10	−0.27	−0.27	I.F.[d]	−0.08	I.F.[d]	I.F.[d]	I.F.[d]	I.F.[d]
OCC	1	39.09	30.71	0.51	1.00	0.17	I.F.[d]	I.F.[d]	−0.17	I.F.[d]	−0.09	I.F.[d]	I.F.[d]	I.F.[d]	−0.31

Measured variables	PLS components	Variance explained by predictor variables, %	Variance explained by measured variables, %	Coefficient of determination, r^2	SECV	Regression coefficients for predictor variables									
						TSS	TOC	pH	EC	Intensity	Duration	Underlying factors			
												N1[c]	N2[c]	N3[c]	N4[c]	N5[c]
TCT	1	60.93	62.19	0.71	1.00	I.F.[d]	0.08	−0.23	−0.18	0.13	−0.11	I.F.[d]	−0.13	I.F.[d]	I.F.[d]	−0.18
DTT	1	65.69	69.02	0.81	0.84	0.60	I.F.[d]	−0.04	I.F.[d]	I.F.[d]	I.F.[d]	I.F.[d]	I.F.[d]	I.F.[d]	I.F.[d]	−0.08
TRT	1	66.51	63.67	0.80	1.00	0.31	−0.17	I.F.[d]	I.F.[d]	I.F.[d]	I.F.[d]	0.37	I.F.[d]	I.F.[d]	I.F.[d]	−0.16
HXT	1	69.56	64.51	0.78	1.00	0.17	I.F.[d]	0.25	I.F.[d]	I.F.[d]	−0.22	−0.07	I.F.[d]	I.F.[d]	0.17	I.F.[d]
OTT	1	73.96	76.76	0.80	0.94	0.27	I.F.[d]	−0.16	I.F.[d]	I.F.[d]	−0.16	0.30	I.F.[d]	I.F.[d]	I.F.[d]	I.F.[d]
TTT	1	65.79	66.04	0.82	0.97	0.23	I.F.[d]	−0.05	I.F.[d]	I.F.[d]	−0.23	0.19	I.F.[d]	0.16	0.15	I.F.[d]

[a] Underlying factors in the light SVOC matrix.
[b] Underlying factors in the heavy SVOC matrix.
[c] Underlying factors in the NVOC matrix.
[d] Insignificant factors in the PLS prediction model for corresponding measured variables.

- SECV ≤ 1;
- 10% maximum difference between the percentage variance explained by the predictor and the measured variables;
- Additional PLS components will only be included if the percentage variance explained by the predictor with the inclusion of an additional PLS component increases by more than 10%.

Table 3 gives the outcome of the PLS regression based on the above criteria.

The outcomes of the PLS regression model was optimised with a reduced number of predictor variables in Table 3. Therefore, not all of the predictor variables were required to predict the individual target components in Table 3. As a final step in the model validation, data matrices were constructed from the remaining rain events that were not used in the construction of the calibration matrices. In this study, the validation of the PLS model was performed by comparing the distributions of the box plot statistics of observed and predicted data matrices for the five size fractions. Fig. 3 shows the distributions for >300 μm and <1 μm size fractions. The box plot statistics for the remaining particulate fractions are shown in the supplementary data.

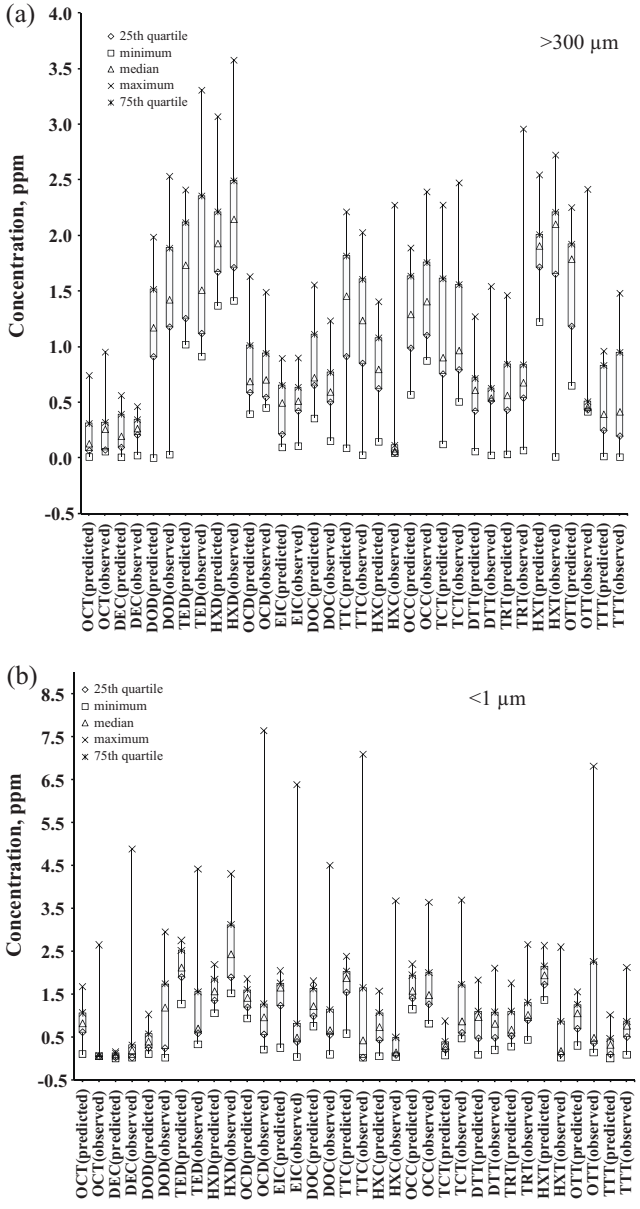

Fig. 3. Distributions of the box plot statistics for observed and predicted target compounds at (a) >300 μm showing similar distributions from minimum to 75th quartile for 14 compounds except DOC, HXC and OTT and at (b) <1 μm showing dissimilar distributions for most compounds.

In Fig. 3a, it is evident that except for DOC, HXC and OTT, the distributions of the concentrations of the remaining 14 target compounds are quite similar from minimum to 75th quartile in the observed and predicted data matrices. Other particulate fractions also showed similar results with very few exceptions. However, the dissolved fraction of <1 μm did not show any such similarity among the box plot statistics in Fig. 3b. This is attributed to the fact that the solubilities of the target compounds in water are very low and these compounds are mainly attached to the particulate solid fractions during their wash-off. In fact, in a very recent study of SVOC and NVOC build-up on urban roads, Mahbub et al. [29] established that SVOCs and NVOCs remain attached primarily with the particulate fractions of 75–300 μm. In order to derive a more comprehensive outlook on the validation of the PLS model, the coefficient of variation (CV%) of the predicted concentrations for the remaining rain events were analysed and the results are shown in Fig. 4.

Horwitz [30] suggested a range of ±20% for the coefficient of variation at the ppm level concentrations estimation of organic compounds in different laboratories. In Fig. 4, the inter-event percentage CV are investigated for the five size fractions and it is clearly evident that the CV% were as high as 55% for the dissolved fraction of <1 μm. However, the CV% were in the range of 5–25% for the particulate fractions from >300 μm to 1 μm with very few exceptions. This also confirmed the fact that the solubility of the target compounds were very low in water and hence, compromises the predictive capacity of the PLS framework for the dissolved fraction of <1 μm. The PLS framework performed with acceptable predictions within the range of minimum to 75th quartile of the observed concentration values at particulate fractions from >300 to 1 μm for the different rainfall characteristics influenced by climate change.

In several chemometric studies where experiments were conducted under stringent laboratory conditions, the experimental design of the calibration matrices gave better prediction results. For example, Sivakumar et al. [31] achieved ≤2% coefficient of variation in an optimisation study aimed at commercial domperidone and pantoprazole preparation with three independent factors assumed significant priori. In another study, Ni et al. [32] reported up to 36% error during the prediction of nitrobenzene and nitro-substituted phenols using the single component PLS method with the number of measured variables taken as significant factors in the data matrices. Both of these studies (i.e. 31 and 32) used chemometric

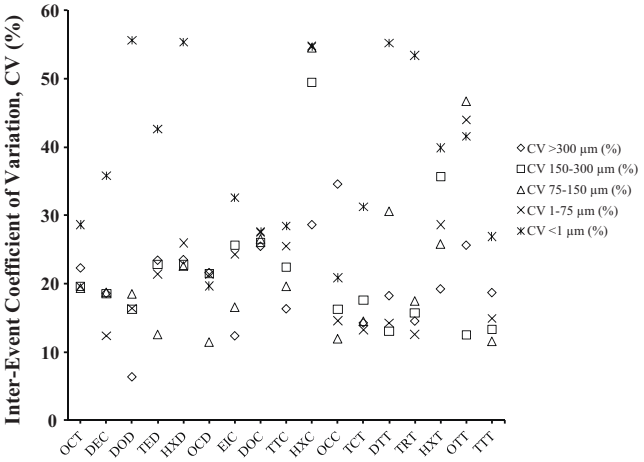

Fig. 4. Coefficient of variations (CV%) of the predicted concentrations at the rain events not used in the calibration.

experimental design to predict the response variables with prior knowledge of independent factors and the calibration and validation sets were prepared under strict laboratory conditions.

The advantage of the current study over the past studies is that this proposed framework allows the introduction of underlying uncorrelated factors into the data matrices. Therefore, it is not necessary to assume significant factors priori. Considering the fact that stringent laboratory conditions could not be applied into the experimental design of the calibration matrices as the wash-off sample collection was field based, the model's ability to predict most of the light SVOCs, heavy SVOCs and NVOCs within an acceptable range provide researchers a robust tool to forecast the concentrations of these pollutants in particulate fractions of wash-off due to climate change induced rainfall characteristics.

4. Conclusions

This research study established a prediction framework for SVOCs and NVOCs under climate change induced rainfall scenario and presented a useful tool for estimating the concentrations of these pollutants under a dynamic situation. The study found that the intensity and durations of low to moderate rain events mainly affect the wash-off of semi- and non-volatile organic compounds from urban roads. The study also proposed that the extraction of the underlying uncorrelated factors within the data matrices constructed from environmental samples can overcome the stringent conditions for the laboratory preparation of calibration and validation matrices for a successful experimental design. The optimisation of the prediction of the wash-off of SVOCs and NVOCs under climate change induced rain events were achieved by considering only the significant variables for a particular compound. The choice of the PLS components based on SECV <1, 10% maximum difference in variances explained by the predictor and measured variables as well as minimum 10% increase in variance by the inclusion of extra components resulted in reduced numbers of significant predictor variables for acceptable prediction performance for the wash-off of SVOCs and NVOCs in particulate fractions >300–1 μm.

Acknowledgements

This study was undertaken as a part of an Australian Research Council funded Linkage Project (LP0882637). The first author gratefully acknowledges the postgraduate scholarship awarded by Queensland University of Technology to conduct his doctoral research. Support from the Gold Coast City Council and Queensland Department of Transport and Main Roads are also gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jhazmat.2012.01.062.

References

[1] Interlaboratory Study of Three Methods for Analysing Petroleum Hydrocarbons in Soils, API Publication No. 4599, American Petroleum Institute, Washington, DC, 1994 (accessed 10.09.10) www.api.org/Publications.
[2] W.M. Draper, J.S. Dhaliwal, S.K. Perera, F.J. Bauman, Determination of diesel fuel and motor oil in water and wastes by a modified diesel-range organics total petroleum hydrocarbon method, J. AOAC Int. 79 (1996) 508–519.
[3] Climate Change in Australia, CSIRO Technical Report 2007, Bureau of Meteorology, Australia, 2007 (accessed 12.09.2010) www.climatechangeinaustralia.gov.au.
[4] L. Bao, S. Chan, L. Wu, T.K. Hei, Y. Wu, Z. Yu, A. Xu, Mutagenicity of diesel exhaust particles mediated by cell–particle interaction in mammalian cells, Toxicology 229 (2007) 91–100.
[5] W.K.C. Morgan, R.B. Reger, D.M. Tucker, Health effects of diesel emissions, Ann. Occup. Hyg. 41 (1997) 643–658.
[6] J.-Y. Ko, J.W. Day, A review of ecological impacts of oil and gas development on coastal ecosystems in the Mississippi Delta, Ocean Coast. Manage. 47 (2004) 597–623.
[7] M. Caselli, G.D. Gennaro, A. Marzocca, L. Trizio, M. Tutino, Assessment of the impact of the vehicular traffic on BTEX concentration in ring roads in urban areas of Bari (Italy), Chemosphere 81 (2010) 306–311.
[8] A.J. Buczynska, A. Krata, M. Stranger, A.F.L. Godoi, V. Kontozova-Deutsch, L. Bencs, I. Naveau, E. Roekens, R.V. Grieken, Atmospheric BTEX-concentrations in an area with intensive street traffic, Atmos. Environ. 43 (2009) 311–318.
[9] P. Wang, W. Zhao, Assessment of ambient volatile organic compounds (VOCs) near major roads in urban Nanjing, China, Atmos. Res. 89 (2008) 289–297.
[10] L. Smith, S. Mukerjee, M. Gonzales, C. Stallings, L. Neas, G. Norris, H. Özkaynak, Use of GIS and ancillary variables to predict volatile organic compound and nitrogen dioxide levels at unmonitored locations, Atmos. Environ. 40 (2006) 3773–3787.
[11] L. Lamon, M. Dalla Valle, A. Critto, A. Macromini, Introducing and integrated climate change perspective in POPs modelling, monitoring and regulation, Environ. Pollut. 157 (2009) 1971–1980.
[12] P. Mahbub, A. Goonetilleke, G.A. Ayoko, Prediction model of the build-up of volatile organic compounds on urban roads, Environ. Sci. Technol. 45 (2011) 4453–4459.
[13] L. Herngren, A. Goonetilleke, R. Sukpum, D.Y. De Silva, Rainfall simulation as a tool for urban water quality research, Environ. Eng. Sci. 22 (2005) 378–383.
[14] D. Abbs, K. McInnes, T. Rafter, The Impact of Climate Change on Extreme Rainfall and Coastal Sea Levels over South-East Queensland, Part 2: A High-Resolution Modelling Study of the Effect of Climate Change on the Intensity of Extreme Rainfall Events, CSIRO Division of Marine and Atmospheric Research, Aspendale, Victoria, Australia, 2007.
[15] Climate change 2007: the physical science basis, in: S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, H.L. Miller (Eds.), Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom/New York, USA, 2007.
[16] P. Mahbub, G. Ayoko, P. Egodawatta, T. Yigitcanlar, A. Goonetilleke, Traffic and climate change impacts on water quality: measuring build-up and wash-off of heavy metals and petroleum hydrocarbons, in: T. Yigitcanlar (Ed.), Rethinking Sustainable Development: Urban management, Engineering and Design, Engineering Science Reference, New York, 2010, pp. 147–167.
[17] P.K. Egodawatta, Translation of small-plot scale pollutant build-up and wash-off measurements to urban catchment scale, Ph.D. Dissertation, Queensland University of Technology, Brisbane, 2007.
[18] J.J. Sansalone, S.G. Buchberger, Partitioning and first flush of metals in urban roadway storm water, J. Environ. Eng. 123 (1997) 134–143.
[19] Test Methods for Evaluating Solid Waste, Physical/Chemical Methods. SW-846, 3rd ed., United States Environmental protection Agency. Methods: 3510C, 8015, 8021, 8260, 2009 (accessed 20.10.09) www.epa.gov/wastes/hazard/testmethods/sw846/.
[20] Method for Determining Diesel Range Organics, Wisconsin DNR 1995, Publication No. PUBL-SW-140, 1995.
[21] J.P.A. Neeft, M. Makkee, J.A. Moulijn, Diesel particulate emission control, Fuel Process. Technol. 47 (1996) 1–69.
[22] G. Andreou, S. Rapsomanikis, Origins of n-alkanes, carbonyl compounds and molecular biomarkers in atmospheric fine and coarse particles of Athens, Greece, Sci. Total Environ. 407 (2009) 5750–5760.
[23] A.P. Kudchadker, B.J. Zwolinski, Vapour pressures and boiling points of normal alkanes, C21 to C100, J. Chem. Eng. Data 11 (1966) 253–255.
[24] A.D. Eaton, L.S. Clesceri, E.W. Rice, A.E. Greenberg (Eds.), Standard Methods for the Examination of Water and Wastewater, 21st ed., American Public Health Association, Washington, DC, 2005.
[25] P. Mahbub, G.A. Ayoko, A. Goonetilleke, P. Egodawatta, S. Kokot, Impacts of traffic and rainfall characteristics on heavy metals build-up and wash-off from urban roads, Environ. Sci. Technol. 44 (2010) 8904–8910.
[26] SPSS, PASW Statistics 18, Release 18.0.1, Tutorials, 2009, http://www.spss.com.
[27] SIRIUS, SIRIUS Version 7.1© copyright 1987–2008, Pattern Recognition Systems AS, Help Topics, 2008, http://www.prs.no.
[28] S. Wold, Cross-validatory estimation of the number of components in factor and principal components models, Technometrics 20 (1978) 397.
[29] P. Mahbub, G.A. Ayoko, A. Goonetilleke, P. Egodawatta, Analysis of the build-up of semi and non volatile organic compounds on urban roads, Water Res. 45 (2011) 2835–2844.
[30] W. Horwitz, Evaluation of analytical methods used for regulation of foods and drugs, Anal. Chem. 54 (1982) 67A–76A.
[31] T. Sivakumar, R. Manavalan, C. Muralidharan, K. Valliappan, Multi-criteria decision making approach and experimental design as chemometric tools to optimize HPLC separation of domperidone and pantoprazole, J. Pharm. Biomed. Anal. 43 (2007) 1842–1848.
[32] Y. Ni, L. Wang, S. Kokot, Simultaneous determination of nitrobenzene and nitro-substituted phenols by differential pulse voltammetry and chemometrics, Anal. Chim. Acta 431 (2001) 101–113.
[33] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics Part A, Elsevier, Amsterdam, 1997.
[34] L.S. Meyers, G. Gamst, A.J. Guarino, Applied Multivariate Research: Design and Interpretation, Sage Publications Inc., California, 2006.

[35] E.M. Peña-Méndez, M.S. Astorga-España, F.J. García-Montelongo, Interpretation of analytical data on n-alkanes and polynuclear aromatic hydrocarbons in *Arbacia lixula* from the coasts of Tenerife (Canary Islands, Spain) by multivariate data analysis, Chemosphere 39 (1999) 2259–2270.
[36] M.S. Yidana, D. Ophori, B. Yakubo, A multivariate statistical analysis of surface water chemistry data—the Ankobra Basin, Ghana, J. Environ. Manage. 86 (2008) 80–87.
[37] X. Zhu, L. Zhang, X. Che, L. Wang, The classification of hydrocarbons with factor analysis and the PONA analysis of gasoline, Chemom. Intell. Lab. Syst. 45 (1999) 147–155.
[38] E.R. Christensen, S. Arora, Source apportionment of PAHs in sediments using factor analysis by time records: application to Lake Michigan, USA, Water Res. 41 (2007) 168–176.
[39] S.N. Deming, S.L. Morgan, Experimental Design: A Chemometric Approach, 2nd ed., Elsevier Science Publishers, Amsterdam, 1993.
[40] C.R.T. Tarley, G. Silveira, W.N.L. dos Santos, G.D. Matos, E.G.P. da Silva, M.A. Bezerra, M. Miró, S.L.C. Ferreira, Chemometric tools in electroanalytical chemistry: methods for optimization based on factorial design and response surface methodology, Microchem. J. 92 (2009) 58–67.
[41] K.R. Beebe, B.R. Kowalski, An introduction to multivariate calibration and analysis, Anal. Chem. 59 (1987) 1007A–1017A.
[42] H. Wang, Q. Liu, Y. Tu, Interpretation of partial least-squares regression models with VARIMAX rotation, Comput. Stat. Data Anal. 48 (2005) 207–219.
[43] Y. Jang, P. Jain, T. Tolaymat, B. Dubey, T. Townsend, Characterization of pollutants in Florida street sweepings for management and reuse, J. Environ. Manage. 91 (2009) 320–327.
[44] A.M. Christensen, F. Nakajima, A. Baun, Toxicity of water and sediment in a small urban river (Store Vejleå, Denmark), Environ. Pollut. 144 (2006) 621–625.